THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial

Citation for published version:
OPPTIMUM study group 2016, 'Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial', *The Lancet*, vol. 387, no. 10033, pp. 2016-2116. https://doi.org/10.1016/S0140-6736(16)00350-0

Digital Object Identifier (DOI):
10.1016/S0140-6736(16)00350-0

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
The Lancet

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 08. May. 2023

Title: A randomised trial of vaginal progesterone prophylaxis for preterm birth: the OPPTIMUM Study

Authors:

Professor Jane Elizabeth Norman MD, Professor Neil Marlow DM, Claudia-Martina Messow PhD, Professor Andrew Shennan, MD, Professor Phillip R Bennett, MD, Professor Steven Thornton DM, Professor Stephen C Robson, MD, Alex McConnachie, PhD, Professor Stavros Petrou, PhD, Professor Neil J Sebire MD, Professor Tina Lavender PhD, S Whyte MSc, Professor J Norrie MSc for the OPPTIMUM study group.

Authors' academic affiliations are listed in the appendix. Address reprint requests to: Jane E Norman at Tommy's Centre for Maternal and Fetal Health, University of Edinburgh MRC Centre for Reproductive Health, Queen's Medical Research Institute, 47 Little France Crescent, Edinburgh

EH16 4TY, UK or jane.norman@ed.ac.uk

Collaborators

Zarko Alfirevic, Liverpool Women's Hospital, Liverpool, UK; Elisabeth Almstrom, Norra Älvsborgs Länssjukhus (NÄL), Trollhättan, Sweden; Ian Arthur, Blackpool Victoria Hospital, Blackpool, UK; Carina Bejlum, Norra Älvsborgs Länssjukhus (NÄL), Trollhättan, Sweden; Rita Arya, Warrington and Halton Hospitals NHS Foundation Trust, Warrington, UK; Karen Brackley, Princess Anne Hospital, Southampton; UK; Helene Brandon, Queen Elizabeth Hospital, Gateshead, UK; Martin Cameron, Norfolk and Norwich University NHS Trust, Norwich, UK; David Churchill, Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK; Fiona Clarke, Burnley General Hospital, Burnley, UK; Janet Cresswell, Chesterfield Royal Hospital, Chesterfield, UK; Lena Crichton, Aberdeen

Maternity Hospital, Aberdeen, UK; Fiona Crosfill, Royal Preston Hospital, Preston, UK; Anna David, University College Hospital, London, UK; Magdy El-Khanagry, Queen's Hospital, Burton-on-Trent, UK; Umo Esen, South Tyneside NHS Foundation Trust, South Shields, UK; Luca Fusi, Ealing Hospital, London North West Healthcare NHS Trust, London, UK; Paul Fogarty, Ulster Hospital, Belfast, UK; Joanna Girling, West Middlesex Hospital, Isleworth, UK; Kath Granger, Morecambe Bay Trust, Lancaster, UK; Henrik Hagberg, Sahlgrenska University Hospital, Gothenburg, Sweden; Elizabeth Haslett, Blackpool Victoria Hospital, Blackpool, UK; Kim Hinshaw, Sunderland Royal Hospital, Sunderland, UK; Sunday Ikhena, Pilgrim Hospital, Boston, UK; Amaju Ikomi, Basildon and Thurrock University Hospital, Basildon, UK; Rabia Imtiaz, Worcester Royal Hospital, Worcester, UK; Bo Jacobssen, Sahlgrenska University Hospital; Gothenburg, Sweden; Tony Kelly, Brighton & Sussex University Hospitals NHS, Brighton, UK; Pihla Kuusela, Sahlgrenska University Hospital; Gothenburg, Sweden; Vanessa Mackay,Queen Elizabeth University Hospital; Glasgow, UK; Shonag Mackenzie, Wansbeck Hospital, Northumberland, UK; Surindra Maharaj, Wishaw General Hospital, Wishaw, UK; Bill Martin, Birmingham Women's Hospital, Birmingham, UK; Elizabeth Martindale, Royal Blackburn Hospital, Blackburn, UK; Basem Muammar, Russells Hall Hospital, Dudley, UK; Stella Mwenechenya, Arrowe Park hospital, Wirral, UK; Shazia Nazir, Pembury Hospital, Maidstone & Tunbridge Wells, UK; Avideah Nejad, Basingstoke & North Hampshire Hospital, Basingstoke, UK; Shaughn O'Brien, North Staffordshire Hospital, Stoke-on-Trent, UK; Odiri Oteri, Lincoln County Hospital, United Lincolnshire Hospitals NHS Trust, Lincoln, UK; Oluseye Oyawoye, Newham University Hospital, London, UK; Shalini Patni, Heart of England Hospital, Birmingham, UK; Donald Peebles, University College Hospital, London, UK; Shanthi Pinto, Leighton Hospital, Crewe, UK; Siobhan Quenby, University Hospital of Coventry, Coventry, UK; Sanaria Raouf, Derby City Hospital, Derby, UK; Alex Rees, University Hospital Wales, Cardiff, UK; Sanjeev Sharma, Southport and Ormskirk

Hospital, Southport, UK; Olanrewaju Sorinola, South Warwickshire NHS Foundation Trust, Warwick, UK; Sarah Stock, Co-Investigator, Royal Infirmary of Edinburgh, Edinburgh, UK; Andrew Tapp, Royal Shrewsbury Hospital, Shrewsbury, UK; Myles Taylor, Royal Devon & Exeter Hospitals NHS Trust, Exeter, UK; Tg Teoh, St Mary's Hospital, London (with Imperial), London, UK; Vasso Terzidou, Chelsea & Westminster Hospital, London, UK; Basky Thilaganathan, St George's Hospital, University of London, London, UK; Jim Thornton, Nottingham City Hospital/ Queen's Medical Centre, Nottingham, UK; A. John Tomlinson, Royal Bolton Hospital, Bolton, UK; Clare Tower, Manchester Academic Science Centre, Manchester, UK; Derek Tuffnell, Bradford Royal Infirmary, Bradford, UK; Graham Tydeman, Forth Park Hospital, Kirkcaldy, Fife, UK; Patrick Vandekerckhove, St Mary's Hospital, Newport - Isle of Wight, UK; Karen Watkins, Royal Cornwall Hospital, Truro, UK; Ross Welch, Derriford Hospital, Plymouth, UK; Ulla-Britt Wennerholm, Sahlgrenska University Hospital, Gothenburg, Sweden.

Background

Progesterone administration has been shown to reduce the risk of preterm birth and neonatal morbidity in women at high risk, but there is uncertainty about longer term effects on the child.

Methods

We conducted a double blind randomised placebo controlled trial of vaginal progesterone, 200mg daily taken from 22-24 to 34 weeks of gestation, on pregnancy and infant outcomes in women at risk of preterm birth (because of previous spontaneous birth at ≤34+0 weeks of gestation, or a cervical length ≤25mm, or because of a positive fetal fibronectin test combined with other clinical risk factors for preterm birth [any one of a history in a previous pregnancy of preterm birth, second trimester loss, preterm premature fetal membrane rupture, or a history of a cervical procedure to treat abnormal smears]). The objective of the study was to determine whether vaginal progesterone prophylaxis, given to reduce the risk of preterm birth, affects neonatal and childhood outcomes. We defined three primary outcomes: (obstetric) fetal death or birth before 34 weeks, (neonatal) a composite of death, brain injury or bronchopulmonary dysplasia, and (childhood) a standardised cognitive score at 2 years of age, imputing values for deaths. Randomisation was carried out through a web portal, with participants, investigators and others involved in giving the intervention, assessing outcomes and/or analysing data remaining masked to treatment allocation until the end of the study. Analysis was by intention to treat. The study was registered and has now closed: ISCRTN number 14568373.

Findings

We recruited 1228 women between 2 February 2009 and 12 April 2013, 610 to the placebo and 618 to the progesterone group. In the placebo group, 597,587 and 439

women or babies were available for analysis of obstetric, neonatal and childhood outcomes respectively; corresponding figures for the progesterone group were 600, 589 and 430. After correction for multiple outcomes, progesterone had no significant effect on the primary obstetric or neonatal outcome: odds ratios OR (95% CI, adjusted for multiple comparisons) of 0.86 (0.61, 1.22) and 0.62 (0.38, 1.03) respectively; nor on the childhood outcome - mean ± SD cognitive scores of 97.3 ± 17.9 (progesterone) and 97.7 ± 17.5 (placebo), difference in means (95% CI) of 0.48, (-2.77, 1.81). There were 70/610 (11.5%) maternal or child serious adverse events in the placebo group and 59/616 (9.8%) in the progesterone group (p=0.27).

Interpretation

In this large study, vaginal progesterone was not associated with reduced risk of preterm birth or composite neonatal adverse outcomes, and had no long term benefit or harm on outcomes in children at two years of age.

Funding

This project was funded by the Efficacy and Mechanism Evaluation (EME) Programme, an MRC and NIHR partnership. The EME Programme is funded by the MRC and NIHR, with contributions from the CSO in Scotland and NISCHR in Wales.

Word count abstract 463

Introduction

Several studies have evaluated either vaginal progesterone or intramuscular 17α-hydroxyprogesterone caproate for the prevention of preterm birth in asymptomatic women with singleton pregnancy at high risk of preterm birth. An individual patient data meta-analysis of women with a short cervix, showed that vaginal progesterone reduced the risk of preterm birth before 33 weeks (relative risk [RR] 0.58; 95% confidence interval [(CI] 0.42–0.80), and reduced a composite of neonatal mortality and morbidity (RR 0.57; 95% CI 0.40-0.81) [1]. Although there is debate whether vaginal and intramuscular therapies have similar mechanisms or efficacy, the Cochrane Library meta-analysis groups the two treatments together, but reports separately for different maternal risk groups. [2] Reduced risk of preterm birth before 34 weeks was demonstrated in women with a short cervix (RR 0.64, 95% CI 0.45 to 0.90), without impact on perinatal mortality or neonatal death: perinatal mortality RR 0.74 (0.42 to 1.29), neonatal death RR 0.55 (0.26, 1.13) [2]. In contrast, in women with previous preterm birth, progestogens reduced the incidence of preterm birth (RR 0.31, 95% CI 0.14, 0.69), perinatal mortality and neonatal death [2]. Notably, although intramuscular 17α-hydroxyprogesterone caproate is licensed for women with a previous preterm birth, a independent analysis of data on vaginal progesterone for a Food and Drug Administration advisory panel showed no benefit, with the panel concluding that “the overall risk/benefit profile *[is]* not acceptable” to support approval of vaginal progesterone in women with a short cervix[3].

Despite recommendations for progesterone use [4] there are few data on long-term benefit or safety for the baby beyond the neonatal period. Adverse childhood effects of preterm birth include neurodevelopmental and cognitive impairments, and increase with degree of prematurity [5]. Progesterone, by delaying birth and reducing prematurity, may reduce risk of

impairment, but this could be offset by direct fetal harm by continuing prolonged exposure to intrauterine infection or inflammation, commonly associated with preterm labour. Furthermore, therapies applied in pregnancy may have differing effects in the neonatal period and early childhood (benefit in one and harm in another), as shown in the ORACLE II trial of antibiotics in spontaneous preterm labour[6,7] and in trials of multiple doses of corticosteroids [8]. Hence, further information on childhood outcomes following progesterone treatment is required to determine the risk/benefit ratio of this therapy.

We therefore conducted a double blind randomised trial to determine whether vaginal progesterone prophylaxis, given to reduce the risk of preterm birth, affects neonatal and childhood outcomes.

Methods

Study design

OPPTIMUM (Does progesterone prophylaxis to prevent preterm labour improve outcome?) is a multicentre randomised double blind placebo controlled trial (ISRCTN: 14568373). The study was granted approval by the Scotland A Research Ethics Committee (Reference 08/MRE00/6). Clinical trials authorisation was given by the Medicines and Healthcare products Regulatory Agency (MHRA reference 22931/0009/001-0001 later revised to 01384/0208/001). An abbreviated protocol has been published [9]. Women were recruited from 65 UK NHS hospitals and one Swedish hospital, from 2 February 2009 until 12 April 2013. Final patient outcome data were collected on 28 August 2015.

Study population

The study comprised a screening phase between 18-24 weeks and 0 days of gestation and a treatment phase, starting at between 22 and 24 weeks. Written informed consent was obtained for both screening and treatment phases at 18-24 weeks and 0 days of gestation and 22 and 24 weeks gestational respectively. All women had a singleton pregnancy, with gestational age established by ultrasound scan before 16 weeks, and were 16 years or older at screening. Women with clinical risk factors for preterm birth (any of a history in a previous pregnancy of preterm birth, or second trimester loss, or preterm premature fetal membrane rupture, or any history of a cervical procedure to treat abnormal smears) and a positive fetal fibronectin test at 22-24 weeks of gestation were eligible for randomisation in the treatment phase from the beginning of the trial, and designated "fibronectin positive". Following analysis of preliminary (blinded) data in July 2010, and the publication of a systematic review on screening for preterm birth [10] we realised that our initial selection strategy erroneously missed women at medium to high risk of preterm birth. Thus from September 2010, after recruitment of the initial 84 women, fibronectin-negative women with (i) a history of spontaneous preterm birth at ≤34 weeks of gestation, or (ii) a cervical length ≤25mm were also eligible for inclusion, and designated a "fibronectin negative" group (see Supplementary Figure 1 for detailed inclusion and exclusion criteria and diagram of fibronectin positive/negative group allocation). Note, there are no nationally agreed recommendations on which pregnant women should be screened for preterm birth risk by measuring cervical length, nor did the OPPTIMUM protocol include recommendations on who should undergo cervical length screening hence any such measurements were made by clinicians on an individual patient basis prior to the woman's recruitment to OPPTIMUM. A cervical length of ≤25mm at any time between 18+0 and 24+0 weeks gestation in the index pregnancy conferred eligibility for recruitment.

Trial intervention

Eligible women were allocated in a 1:1 ratio to either progesterone 200mg soft capsules (Utrogestan, Besins Healthcare) or an identical appearing placebo. The participant administered the vaginal study medication daily at bedtime, commencing from around (22 to 24 weeks of gestation) until 34 weeks or delivery of the baby, whichever was sooner. Co-administration of bromocriptine, rifamycin, ketoconazole or ciclosporin was prohibited due to potential drug interactions. Rules for individual women to stop treatment on safety grounds (e.g. after development of symptomatic placenta praevia) are defined in the protocol.

Randomisation and masking

Assignment to treatment allocation was carried out through a web portal hosted by the study data centre at the Robertson Centre for Biostatistics, at the Glasgow Clinical Trials Unit, University of Glasgow. The randomisation schedule was computer-generated at the Robertson Centre, using the method of randomised permuted blocks of length four, stratified by history of a previous pregnancy of more than 14 weeks of gestation and by study centre. Allocation concealment was achieved by use of a placebo, which appeared identical to the active drug. Participants were asked for informed consent and enrolled by collaborating clinicians (listed above and in the supplementary), who used the web portal described above to randomise participants to treatment. Treatment allocation corresponded to a box number in the local pharmacy, containing either active or placebo drug. Participants, investigators, pharmacists and others involved in giving the intervention, assessing outcomes and/or analysing data remained masked to treatment allocation until the end of the study. There was no formal attempt made to evaluate the success of masking.

Compliance (assessed for each woman using a combination of medication packs returns, patient diaries and patient self reports) was calculated as the percentage of doses of study medication used divided by the expected doses. Adequate compliance was taken as 80% of prescribed medication.

Outcome ascertainment

Participants and investigators were masked to treatment allocation until all outcomes were assessed and the study database locked. Data were collected at screening, randomisation, 34 weeks of gestation, during labour and delivery, during the neonatal stay and at one and two years post delivery to determine clinical outcomes. Two-year evaluations, based on chronological age because of the mixed term/preterm population, were undertaken at the local hospital clinic or at home. This comprised the parent-completed structured clinical history, a parent-completed behavioral measure (the "Strengths and Difficulties Questionnaire") and the Cognitive scale of the Bayley Scales of Infant and Toddler Development 3rd Edition (Bayley-III). All evaluations were undertaken by assessors who had received training, either from the study centre or via a national course; all met pre-specified criteria of 90% agreement or more on an item-by-item basis with an independent psychologist. Record forms were checked centrally for consistency and completeness. For children for whom we could not arrange a clinic assessment we requested information from the family doctor concerning general health and the presence of motor, sensory and developmental concerns. Outcomes were categorised as moderate or severe using published definitions [11].

We defined three primary outcomes:

1. **obstetric**: fetal death or delivery either occurring before 34+0 weeks of gestation;
2. **neonatal**: a composite of death, bronchopulmonary dysplasia and brain injury on cerebral ultrasound;

3. **childhood**: the Bayley-III cognitive composite score at 22-26 months of chronological age.

Brain injury was defined as any intraventricular hemorrhage (excluding subependymal hemorrhages), parenchymal cystic or hemorrhagic lesion, or persistent ventriculomegaly (ventricular index >97th percentile). All scans were reported locally. All abnormal scans and 10% of normal scans were reviewed centrally masked to the local report (NM). Bronchopulmonary dysplasia (severe chronic lung disease) was defined as need for ≥30% oxygen and/or positive pressure (positive pressure ventilation or nasal continuous positive airway pressure) at 36 weeks postmenstrual age or discharge, whichever came first. A variety of secondary efficacy and safety outcomes were collected as defined in the protocol [9].

Statistical analysis

A statistical analysis plan was finalised prior to data lock. Statistical analyses were performed by C-MM and AMcC at the Robertson Centre for Biostatistics, Glasgow University according to intention to treat. The three primary outcomes and secondary outcomes were compared between the treatment groups using mixed effects logistic regression (or, for continuous variables, linear regression) models including treatment allocation and previous pregnancy (≥ 14 weeks) as fixed effects, with study centre as a random effect. According to the pre-specified statistical analysis plan, p-values were initially reported without adjustment for multiple comparisons, then adjusted using a Bonferroni-Holm procedure [12]. The planned sample size was around 1125 participants, depending on the relative numbers of fFN positive and fFN negative women recruited[9]. Detailed sample size calculations are available in the published protocol, but in brief the

study had at least 80% power to detect what was considered the minimal important clinical difference for each of the three primary outcomes at a nominal 5% level of significance [9].

Sensitivity analyses included repeating the primary analyses in a per-protocol dataset (which excluded data from women who were found not to be compliant with the inclusion / exclusion criteria, or who had a structural or chromosomal fetal anomaly discovered after inclusion, or who had a multiple pregnancy discovered after inclusion or who were not adequately compliant with treatment by the pre-specified definition described above), and the use of multiple imputation of missing primary outcome data. Preplanned subgroup analyses for primary outcomes were performed by extending the main regression models to include interaction terms for the following subgroups: fFN positive/negative, cervical length ≤ 25mm/> 25mm, cervical length ≤ 15mm/>15mm, chorioamnionitis yes/no, history of spontaneous preterm birth/no such history, history of preterm birth / no such history. Safety outcomes (adverse events) were assessed in a “safety” population, excluding women for whom it was documented that no study medication was taken.

Study oversight

A Trial Steering Committee and a Data Monitoring Committee supervised the conduct of the study (Supplementary text 1).

Role of the funding source.

Neither the funders of the study, nor the provider of active and placebo medication had any role in study design, data collection, data analysis, data interpretation, or writing of the report. C-MM and AMcC had full access to all the data in the study and JEN had final responsibility for the decision to submit for publication.

Results

Participation and compliance

We screened 15,132 women: 1,228 (8%) were subsequently entered into the study, 610 allocated to placebo and 618 to progesterone (Figure 1). Two of these women were randomised in error and were excluded from initiating on treatment and the intention to treat (ITT) population. Baseline characteristics of participants in the ITT population were balanced across the two allocated groups (Table 1). The number of women randomised per site ranged from 1-165; three sites screened but did not randomise participants. Information on the obstetric, neonatal and childhood primary outcomes was available for 1,197 (97.4%), 1,176 (95.8%) and 869 (70.8%) of participants respectively. There were few differences in baseline characteristics between those for whom primary outcome data was or was not available (Supplementary Table 1).

Information available from diary returns for 1,011 (82.3%) women, showed ≥80% compliance in 361/509 (71%) in the placebo group and 333/502 (66.3%) in the progesterone group. For compliant women, the median percentage of medication taken was 92.3 (quartiles: 71.6-98.7) and 92.9 (quartiles: 59.0-98.6), respectively. No woman terminated treatment because of pre-specified discontinuation rules.

Primary outcomes

Although the point estimate of the OR was in the direction of benefit, administration of progesterone did not significantly alter the risk of the obstetric (fetal death or birth before 34 weeks) or neonatal (a composite of death, brain injury or bronchopulmonary dysplasia) outcome after the pre-specified adjustment for multiple comparisons (Bonferroni-Holm procedure) OR (95% adjusted CI) of 0.86 (0.61, 1.22); and 0.62 (0.38, 1.03) respectively (Table 2). Similarly there was no effect on childhood outcomes: mean ± SD cognitive score

of 97.7 ±17.5 (placebo) and 97.3 ± 17.9 (progesterone), difference (adjusted 95% CI) in means of -0.48 (-2.77, 1.81).

Components of the primary outcome

Amongst the components of the primary obstetric and neonatal outcomes, the proportion of babies with observed neonatal brain injuries on cerebral ultrasound scanning was lower in the progesterone arm (3.1% v 5.9%; OR: 0.50; 95%CI: 0.31, 0.84) (Table 2). A reduction in brain injury was also observed in a sensitivity analysis restricted to participants in whom a neonatal brain scan was performed (n=776) (OR 0.54, 95% CI 0.32, 0.88). Neonatal death was also less common in the progesterone group, but the low numbers precluded planned adjustment for the covariate 'previous pregnancy ≥ 14 weeks' gestation.

Similar results for primary outcomes were achieved in per protocol analyses (56.0% of the ITT population) (Supplementary Table 2); in analyses with multiple imputations of missing data on the primary outcomes (Supplementary Table 3); and in alternative multiple comparison procedures, including the Sidak-Holm method and permutation adjustment (50000 permutations) (data not shown). Comparison of characteristics of women included and not included in the per-protocol analysis are shown in Supplementary Table 4. An additional sensitivity analysis with imputations for the variable "smoking" was performed post hoc because of the difference in smoking prevalence in those with and without outcome data: again this generated similar results to the main analysis (data not shown). A post hoc survival curve of time to death or delivery (primary obstetric outcome) showed that the differences between the progesterone and placebo groups appeared greatest at our pre-specified gestational cut off of 34 weeks (Supplementary Figure 2).

Subgroup analyses

Rates of preterm birth were higher in the predefined subgroups of women with a positive fetal fibronectin test, women with a cervical length ≤25mm and women with a cervical length ≤15mm (Supplementary Table 5). However, there were no significant interactions between these groups and the effect of progesterone on any of the obstetric, neonatal or childhood outcomes. Within subgroups there was no significant effect of progesterone on any of the primary outcomes (Table 3). The interaction term approached statistical significance (p=0.05) for the neonatal outcome in the subgroup with a history of a previous spontaneous preterm birth, where the OR (95% CI) for the neonatal outcome was lower in the progesterone group (0.49, 0.30, 0.80); compared with 1.20 (0.56, 2.59) in the complementary group with no previous spontaneous preterm birth. However, caution is needed in interpreting all these findings given the number of pre-specified subgroup analyses undertaken on three primary outcomes.

Other Secondary outcomes

The proportions of other secondary outcomes largely did not differ statistically between progesterone and placebo (Table 4). Although neurodevelopmental impairments were similarly distributed in each group, somatic impairments in renal, gastrointestinal and respiratory systems though of low frequency, were more common in the progesterone group. There were no apparent differences in the proportions with safety or other outcomes between the placebo and progesterone groups (Table 5).

Discussion

OPPTIMUM is the largest randomised trial of vaginal progesterone for prevention of preterm birth in women at high risk. In contrast to published reports [13-15] we show no effect of progesterone on rates of either preterm birth or neonatal composite outcome. For the

first time using a direct assessment, we provide strong evidence that the use of progesterone from 24-34 gestational weeks has no demonstrable effect on 2-year neurodevelopmental outcomes, either as cognitive scores or impairments, suggesting that progesterone prophylaxis to prevent preterm birth appears safe for the baby (at least up to two years of age). Only one previous study has determined long term effects of progestogens given to singleton pregnancies in a randomised trial of intramuscular 17-α hydroxyprogesterone caproate [16,17], but this study used parent report and had a smaller sample size with a higher proportion lost to follow up. The other published studies are limited to questionnaire or health record-based assessments in twins whose mothers were enrolled in randomised trials of progesterone versus placebo [18,19].

OPPTIMUM was a pragmatic trial, set up to examine effects of progesterone on outcomes in a heterogenous group of women at risk of preterm birth. As described in the methods section, we extended our recruitment criteria early in the study, when newly available information suggested we were missing women at high risk of preterm birth. Notably, the “fibronectin negative” group (recruited under the extended criteria), had rates of the primary outcome (death or preterm birth before 34 weeks) of 12.6% (Supplementary Table 4), some three times those of the background population of pregnant women in the UK [20]. Hence our decision to extend the recruitment criteria appears correct. Importantly, although we were able to define at baseline, subgroups of women with higher rates of preterm birth (including those with a short cervix and those with a positive fibronectin test), our data suggest that the efficacy of progesterone (for all outcomes) is similar across groups. Therefore our data do not support the premise that vaginal progesterone is specifically effective in women with a short cervix.

Although we showed no overall effect, point estimates of the reduction in the odds of the obstetric outcome (0.86) and the neonatal composite outcome (0.62) are in the direction of benefit, but with confidence intervals that include no advantage. Additionally, point estimates in the short cervix subgroups are similar to those reported in meta-analyses of the effect of progesterone in such women. For example, the OR for preterm birth prevention was 0.69 in OPPTIMUM, compared with a RR of 0.64 [before 34 weeks] in one systematic review [21] and RR of 0.58 [before 33 weeks] in an individual patient data meta-analysis [1]. The corresponding figures for effects on a neonatal composite are 0.54 in OPPTIMUM (OR) and RR 0.57 in the individual patient data meta-analysis [1]. An individual patient data meta-analysis, including the OPPTIMUM findings, to understand what the totality of evidence indicates, particularly within subgroups of interest, is likely to be helpful.

Although we have shown no significant effect on the overall neonatal composite outcome, there appeared to be a reduction in neonatal brain injury. Progesterone associated reduction in brain injury is plausible given supportive pre-clinical data in adult models showing potentially neuro-protective effects of reduced inflammatory cytokine production, reduced activation of microglial cells and limited apoptosis [22,23], although a recent trial of over 1,000 adult participants with traumatic brain injury has shown no clinical therapeutic effect [24]. However, in the absence of long term improvements in cognitive function, a protective effect of progesterone on brain injury (defined by ultrasound) may not be important clinically: not only was brain injury on ultrasound a relatively rare event in OPPTIMUM but other studies have shown no correlation between this finding and longer term neurosensory impairment [25]. Additionally, these (non-significant) reductions in the neonatal composite adverse outcome need to be considered against the (non-significant) increase in the childhood adverse outcome of death or moderate/severe neurodevelopmental impairment.

OPPTIMUM strongly suggests that the efficacy of progesterone in improving outcomes is either non-existent or weak. Given the heterogeneity of the “preterm labour syndrome” we cannot exclude benefit in specific phenotypic or genotypic subgroups of women at risk. However, the subgroups of women who may benefit do not appear to be easily identifiable by current selection strategies (including cervical length measurement and fibronectin testing).

Reassuringly, our study suggests that progesterone is “safe” for those who wish to take it for preterm birth prophylaxis. The overall rate of maternal or child adverse events was similar in the progesterone or placebo groups. There were few differences in the incidence of adverse secondary outcomes in the two groups, with the exception of a higher rate of renal, gastrointestinal and respiratory complications in childhood in the progesterone groups. Importantly, the absolute rates of these complications was low. Follow up of other babies exposed *in utero* to vaginal progesterone would be helpful in determining whether the increased rate of some renal, gastrointestinal and respiratory complications is a real effect, or Type I error.

A potential weakness in our trial is that overall compliance was only 68.6%. This contrasts with a reported compliance of 88.5% in the Hassan study [13], but is greater than the compliance seen in routine clinical practice [26]. Additionally, the assumption in the Hassan study that women who did not return study medication were fully compliant might have erroneously inflated their estimate of compliance. No information on compliance was reported in the other large study on vaginal progesterone in singletons [15]. Notably, in OPPTIMUM, the magnitude of the effect size for each of the primary outcomes was very similar in the per protocol analysis (restricted to those with adequate treatment

compliance) compared to the intention to treat group; suggesting that suboptimal compliance did not have a major impact on overall results.

We believe that OPPTIMUM should prompt a major review of the use of progesterone for preterm birth prophylaxis, a search to identify specific women who might specifically benefit, and a redoubling of efforts to find alternative strategies to prevent preterm birth in women at risk. For those clinicians and women who wish still to use progesterone for preterm birth prophylaxis, our data provide reassurance that it appears safe, at least until 2 years of age of the child.

Word count main paper 3627

References

1. Romero R, Nicolaides K, Conde-Agudelo A, et al. Vaginal progesterone in women with an asymptomatic sonographic short cervix in the midtrimester decreases preterm delivery and neonatal morbidity: a systematic review and metaanalysis of individual patient data. Am J Obstet Gynecol 2012; 206(2): 124 e1-19.

2. Dodd JM, Jones L, Flenady V, Cincotta R, Crowther CA. Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth. Cochrane database of systematic reviews (Online) 2013; 7: CD004947.

3. Soule S. Meeting of the Advisory Committee for Reproductive Health Drugs, January 20, 2012, 2012.

4. Progesterone and preterm birth prevention: translating clinical trials data into clinical practice. Am J Obstet Gynecol 2012; 206(5): 376-86.

5. MacKay DF, Smith GC, Dobbie R, Pell JP. Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. PLoS Med 2010; 7(6): e1000289.

6. Kenyon S, Pike K, Jones DR, et al. Childhood outcomes after prescription of antibiotics to pregnant women with spontaneous preterm labour: 7-year follow-up of the ORACLE II trial. Lancet 2008; 372(9646): 1319-27.

7. Kenyon SL, Taylor DJ, Tarnow-Mordi W. Broad-spectrum antibiotics for spontaneous preterm labour: the ORACLE II randomised trial. ORACLE Collaborative Group. Lancet 2001; 357(9261): 989-94.

8. Wapner RJ, Sorokin Y, Mele L, et al. Long-term outcomes after repeat doses of antenatal corticosteroids. N Engl J Med 2007; 357(12): 1190-8.

9. Norman JE, Shennan A, Bennett P, et al. Trial protocol OPPTIMUM- Does progesterone prophylaxis for the prevention of preterm labour improve outcome? BMC Preg Childbirth 2012; 12(1): 79.

10. Honest H, Forbes CA, Duree KH, et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. Health Technol Assess 2009; 13(43): 1-627.

11. Marlow N. Measuring neurodevelopmental outcome in neonatal trials: a continuing and increasing challenge. Archives of disease in childhood 2013; 98(6): F554-8.

12. Holm S. A simple sequentially rejective multiple test procedure. Scan J Stats 1979; 6: 65-70.

13. Hassan SS, Romero R, Vidyadhari D, et al. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. Ultrasound Obstet Gynecol 2011.

14. da Fonseca EB, Bittar RE, Carvalho MH, Zugaib M. Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: a randomized placebo-controlled double-blind study. Am J Obstet Gynecol 2003; 188(2): 419-24.

15. Fonseca EB, Celik E, Parra M, et al. Progesterone and the risk of preterm birth among women with a short cervix. N Engl J Med 2007; 357(5): 462-9.

16. Northen AT, Norman GS, Anderson K, et al. Follow-up of children exposed in utero to 17 alpha-hydroxyprogesterone caproate compared with placebo. Obstet Gynecol 2007; 110(4): 865-72.

17. Meis PJ, Klebanoff M, Thom E, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. N Engl J Med 2003; 348(24): 2379-85.

18. McNamara HC, Wood R, Chalmers J, et al. STOPPIT Baby Follow-up Study: the effect of prophylactic progesterone in twin pregnancy on childhood outcome. PLoS One 2015; 10(4): e0122341.

19. Rode L, Klein K, Nicolaides K, Krampl-Bettelheim E, Tabor A. Prevention of preterm delivery in twin gestations (PREDICT): A multicentre randomised placebo-controlled trial on the effect of vaginal micronised progesterone. Ultrasound Obstet Gynecol 2011.

20. Statistics OfN. Birth Characteristics in England and Wales 2014. 2015. http://www.ons.gov.uk/ons/rel/vsob1/birth-characteristics-in-england-and-wales/2014/stb-birth-characteristics-2014.html (accessed 25.01.2015.

21. Dodd JM, Flenady VJ, Cincotta R, Crowther CA. Progesterone for the prevention of preterm birth: a systematic review. Obstet Gynecol 2008; 112(1): 127-34.

22. Stein DG. Progesterone in the treatment of acute traumatic brain injury: a clinical perspective and update. Neuroscience 2011; 191: 101-6.

23. Deutsch ER, Espinoza TR, Atif F, Woodall E, Kaylor J, Wright DW. Progesterone's role in neuroprotection, a review of the evidence. Brain Res 2013; 1530: 82-105.

24. Skolnick BE, Maas AI, Narayan RK, et al. A clinical trial of progesterone for severe traumatic brain injury. N Engl J Med 2014; 371(26): 2467-76.

25. Schmidt B, Davis P, Moddemann D, et al. Long-term effects of indomethacin prophylaxis in extremely-low-birth-weight infants. N Engl J Med 2001; 344(26): 1966-72.

26. Cutler DM, Everett W. Thinking outside the pillbox--medication adherence as a priority for health care reform. N Engl J Med 2010; 362(17): 1553-5.

Acknowledgements

We thank the members of the Trial Steering and Data Monitoring Committee and all the people who helped in the conduct of the study (including the OPPTIMUM collaborative group and other clinicians listed in the appendix). We are grateful to Paul Piette (Besins Healthcare Corporate, Brussels, Belgium) and Besins Healthcare for their kind donation of active and placebo drug for use in the study, and to staff of the pharmacy departments and R&D departments of the participating hospitals. We are also grateful to the many people who helped in this study but who we have been unable to name, and in particular all the women (and their babies) who participated in OPPTIMUM.

OPPTIMUM was funded by the Efficacy and Mechanism Evaluation (EME) Programme, an MRC and NIHR partnership, award number G0700452, revised to 09/800/27. The EME Programme is funded by the MRC and NIHR, with contributions from the CSO in Scotland and NISCHR in Wales. The views expressed in this publication are those of the author(s) and not necessarily those of the MRC, NHS, NIHR or the Department of Health. The funder had no involvement in data collection, analysis or interpretation, and no role in the writing of this manuscript or the decision to submit for publication.

Declarations of interest

JEN reports grants from Medical Research Council, non-financial support from Besins (donation of drugs for OPPTIMUM) during the conduct of the study; grants from other government bodies, including National Institute for Health Research, grants from Tommy's the baby charity, and activity on a Data Safety and Monitoring Committee for Glaxo Smith Kline, outside the submitted work. JEN has also chaired the recent National Institute for Clinical Excellence (NICE - UK) guideline development group on Preterm Labour and Birth. The NICE guidelines were finalised before the OPPTIMUM data were available. NM reports personal fees from Shire, personal fees from Novartis, outside the submitted work

and funding from the Department of Health's NIHR Biomedical Research Research Centres funding scheme at UCLH/UCL. AMcC reports grants from MRC / NIHR EME, during the conduct of the study. C-M M reports grants from MRC / NIHR EME, during the conduct of the study. PRB reports personal fees and other from ObsEva Pharmaceuticals, personal fees from GSK, outside the submitted work and funding from the Department of Health's NIHR Biomedical Research Research Centres funding scheme at Imperial. In addition, PRB has a patent microRNA markers to predict cervical shortening and preterm birth issued. AS reports other from Honoraria and travel expenses for lectures by Hologic, outside the submitted work. ST reports grants from MRC,during the conduct of the study; personal fees from GSK, personal fees and non-financial support from Ferring Pharmaceuticals, grants from Hologic, outside the submitted work.

Contributors

JEN, NM, AS, PRB, ST, SCR, SP, NJS, TL and JN contributed to study design, JEN, NM, AS, PRB, ST, SCR, SP, NJS, TL, SW and the OPPTIMUM study group contributed to data collection, C-M M, A McC and JN performed the statistical analysis of the data, JEN, C-MM, AMcC and JN performed the initial data interpretation, JEN wrote the first draft of the manuscript. All authors contributed to final data interpretation and contributed to and approved the final draft of the manuscript.

Table 1 – Demographics and baseline characteristics of women entered into the treatment phase of the OPPTIMUM study

Characteristic	Placebo (N=610)	Progesterone (N=616)
General		
Age - years	31.4 ± 5.8	31.5 ± 5.6
Smoking – N (%)	125 (20.6)	111 (18.1)
Alcohol- no (%)	34 (5.6)	29 (4.7)
Drug use - no (%)	9 (1.5)	8 (1.3)
Years in full-time education[a]	13.5 (3.0)	13.5 (3.1)
Ethnic group		
- White	446 (73.2)	449 (73.0)
- Black	95 (15.6)	85 (13.8)
- Asian	51 (8.4)	53 (8.6)
- Mixed	12 (2.0)	16 (2.6)
- Other	5 (0.8)	12 (2.0)
Height - cm	163.6 ± 6.4	163.5 ± 6.7
Weight - kg	71.4 ± 16.7	71.9 ± 17.5
BMI (kg/m^2)	26.7 ± 6.1	26.9 ± 6.4
Systolic Blood Pressure - mmHg	112.4 ± 12.2	111.3 ± 12.5
Diastolic Blood Pressure - mmHg	66.2 ± 8.6	65.7 ± 8.5
Obstetric history		
Any previous pregnancy -N (%)	581 (95.4)	591 (96.1)
Previous pregnancy of at least 14 weeks - N (%)	571 (93.8)	578 (94.0)
History of preterm birth (any) - N (%)	473 (78)	493 (80)
Spontaneous preterm birth - N (%)	448 (75)	473 (78)

History of live birth followed by neonatal death - N (%)	85 (14.0)	80 (13.0)
History of stillbirth -N (%)	48 (7.9)	47 (7.6)
Cervical length – mm[b]	28.8 ± 11.1	28.2 ± 10.6
Cervical length ≤25mm[b] - N (%)	119 (33.9)	137 (38.0)
Cervical length ≤15mm [b] - N (%)	47 (13.4)	51 (14.1)
Fibronectin testing in screening phase		
Gestation (weeks) at fFN test	22.9 ± 0.6	22.9 ± 0.6
Positive fFN test result - N (%)	180 (29.7)	163 (26.5)

Continuous variables are summarised as mean ± standard deviation (plus-minus values), categorical variables as number (percentage) per category. All variables are missing for a maximum of 5 women in each group, except where specified otherwise.

[a] Missing for 22 / 31 women in placebo / progesterone group respectively.

[b] Missing for 259 / 255 women in placebo / progesterone group respectively.

Table 2 – Primary outcomes and their components for women entered into the treatment phase of the OPPTIMUM study and their babies.

Primary outcomes	**Placebo**	**Progesterone**	**Odds ratio (95%CI) (unadjusted)**	**P value (unadjusted)**	**Odds ratio (95%CI) (adjusted[a])**	**P value (adjusted[a])**
	Number / denominator (%) or mean ± SD					
Fetal death or delivery < 34 weeks of gestation	108/597 (18.1)	96/600 (16.0)	0.86 (0.64, 1.17)	0.34	0.86 (0.61, 1.22)	0.67
Neonatal morbidity or death	60/587 (10.2)	39/589 (6.6)	0.62 (0.41, 0.94)	0.02	0.62 (0.38, 1.03)	0.072
			Difference in means (95%CI)	**P value (unadjusted)**	**Difference in means (95%CI)**	**P value (adjusted[a])**
Cognitive Composite score at 2 years[b c]	97.7 ± 17.5	97.3 ±17.9	-0.48 (-2.77, 1.81)	0.68	-0.48 (-2.77, 1.81)	0.68

Components of the obstetric outcome	**Placebo**	**Progesterone**	**Odds ratio (95%CI) (unadjusted)**	**P value (unadjusted)**
Fetal death	7/597 (1.2)	8/600 (1.3)	1.14 (0.41, 3.17)	0.8
Liveborn delivery before 34 weeks	101/590 (17.1)	88/592 (14.8)	0.85 (0.62, 1.15)	0.29

Components of the neonatal outcome	Placebo	Progesterone	Odds ratio (95%CI) (unadjusted)	P value (unadjusted)
Neonatal death	6/597 (1.0)	1/600 (0.2)	0.17 (0.06, 0.49)	0.0009[d]
Bronchopulmonary dysplasia[e]	18/574 (3.1)	17/580 (2.9)	0.94 (0.49,1.78)	0.84
Brain injury on ultrasound scan[f]	34/574 (5.9)	18/ 584 (3.1)	0.50 (0.31, 0.84)	0.008

Notes

Binary outcomes are summarised as number of women with outcome / number of women with information on outcome (percentage). Continuous outcomes are summarised as mean ± standard deviation, with n indicating the number with information on outcome.

[a] Confidence interval for odds ratio and p-value adjusted for multiple primary outcomes using Bonferroni-Holm method.

[b] Median [interquartile range] weeks of age at assessment of 111.6 [104.6, 122.2] and 110.4 [104.0, 121.5] in the placebo and progesterone group respectively

[c] Sample size of 439 in the placebo and 430 in the progesterone group respectively

[d] Unadjusted for previous pregnancy of > 14 weeks because of small sample size

[e] Bronchopulmonary dysplasia defined as need for ≥30% oxygen to maintain oxygen saturation above 92% and/or positive pressure (positive pressure ventilation or nasal continuous positive airway pressure) at 36 weeks post menstrual age or discharge, which ever comes first

[f] Brain injury on ultrasound scan defined as any intraventricular hemorrhage (IVH) (excludes subependymal hemorrhages), parenchymal cystic or hemorrhagic lesion or persistent ventriculomegaly (VI >97th percentile). The components of the brain scan abnormalities were: intraventricular hemorrhage - 13/383 (3.4%) and 7/357 (2.0%), parenchymal cystic or hemorrhagic lesion – 23/383 (6.0%) and 8/357 (2.2%), and persistent ventriculomegaly (ventricular index > 97th centile) 8/372 (2.2%) and 3/349 (0.9%) in the placebo and the progesterone groups respectively

Table 3 Pre-specified subgroup analyses based on baseline risk factors in women entered into the treatment phase of the OPPTIMUM study

Primary outcome	**Treatment effect (95% confidence interval), p-value, N**								**Interaction p-value**
	OR	95% CI	p	N	OR	95% CI	p	N	
Fibronectin status	***Negative fFN***				***Positive fFN***				
Obstetric	0.88	(0.58, 1.33)	0.542	859	0.91	(0.57, 1.46)	0.707	338	0.91
Neonatal	0.65	(0.37, 1.13)	0.129	847	0.64	(0.34, 1.20)	0.162	329	0.96
Childhood	-0.63	(-3.28, 2.03)	0.644	628	-1.09	(-5.41, 3.23)	0.612	241	0.86
Cervical length at baseline	***> 25mm***				***≤ 25mm***				
Obstetric	0.88	(0.50, 1.57)	0.672	445	0.69	(0.39, 1.20)	0.191	251	0.54
Neonatal	0.74	(0.35, 1.56)	0.432	436	0.54	(0.25, 1.16)	0.113	246	0.56
Childhood	-2.27	(-6.10, 1.56)	0.247	317	-2.15	(-7.23, 2.93)	0.408	179	0.97
Cervical length at baseline	***> 15mm***				***≤ 15mm***				
Obstetric	0.77	(0.48, 1.23)	0.274	599	0.91	(0.41, 2.04)	0.819	97	0.73
Neonatal	0.73	(0.39, 1.38)	0.334	588	0.49	(0.18, 1.31)	0.156	94	0.50
Childhood	-2.49	(-5.77, 0.78)	0.137	423	-0.69	(-8.60, 7.22)	0.865	73	0.68
Chorioamnionitis	***No***				***Yes***				
Obstetric	1.38	(0.55, 3.45)	0.497	115	2.17	(0.68, 6.85)	0.190	57	0.55

Neonatal	0.81	(0.22, 2.96)	0.752	115	2.21	(0.76, 6.40)	0.148	56	0.24
Childhood	-2.30	(-10.30, 5.70)	p=0.575	81	-1.08	(-11.91, 9.76)	p=0.846	43	0.86
History of spontaneous preterm birth		***No***				***Yes***			
Obstetric	0.99	(0.51, 1.92)	0.972	273	0.82	(0.58, 1.16)	0.254	903	0.62
Neonatal	1.22	(0.55, 2.71)	0.620	270	0.48	(0.29, 0.79)	0.0042	886	0.053
Childhood	-1.11	(-5.96, 3.73)	0.653	201	-0.14	(-2.79, 2.52)	0.919	656	0.73
History of any preterm birth		***No***				***Yes***			
Obstetric	1.06	(0.53, 2.12)	0.868	250	0.81	(0.58, 1.14)	0.225	946	0.50
Neonatal	1.09	0.48, 2.45)	0.836	248	0.52	(0.32, 0.84)	0.0079	927	0.12
Childhood	-0.91	(-5.92, 4.11)	0.724	187	-0.37	(-2.96, 2.23)	0.782	681	0.85

Logistic or linear mixed effects regression models predicting outcome from treatment, subgroup and the interaction of treatment with the subgroup variable, adjusting for previous pregnancy of at least 14 weeks and a random effect for centre. The estimated treatment effect is presented within each subgroup as odds ratio or expected mean difference, as appropriate.

Table 4 - Secondary outcomes

	Placebo	Progesterone	OR, hazard ratio or mean difference (95 % confidence interval)	P value
	Number / denominator (%) or mean ± SD			
Obstetric and neonatal				
Gestational age at delivery (weeks)	36.8 ± 4.2, n=597	36.9 ± 4.1, n=600	1.03 (0.92,1.15)	0.62
Deaths up to 2 years of age	16 / 509 (3.1)	20 / 500 (4.0)	1.28 (0.66, 2.5)	0.47
Death after trial entry up to end of study	16 / 598 (2.7)[a]	20 / 600 (3.3)[a]	1.26 (0.65, 2.42)	0.5
Daily category of care after delivery room (normal / special / high dependency / intensive)				
Number of days of normal care	1.7 ± 2.3, n=570	1.7 ± 1.6, n=581	n/a	n/a
Number of days of special care	4.2 ± 10.6, n=570	2.9 ± 8.3, n=581	n/a	n/a
Number of days of high dependency care	2.2 ± 8.4, n=569	2.1 ± 10.4, n=580	n/a	n/a

Number of days intensive care	1.8 ± 7.3, n=569	1.9 ± 8.1, n=580	n/a	n/a
Surfactant administration	45/573 (7.9)	47/583 (8.1)	1.03 (0.68,1.55)	0.9
Suspected or confirmed necrotising entercolitis	13/574 (2.3)	18/581 (3.1)	1.37 (0.76, 2.45)	0.29
Infections				
Neonatal infection	36/573 (6.3)	44/537 (7.6)	1.22 (0.79,1.88)	0.36
One or more discrete episodes with positive blood culture amongst those with infection	19/33 (57.6)	17/40 (42.5)	0.51 (0.19.1.34)	0.18
One or more discrete episodes with positive CNS culture amongst those with infection	0/34 (0)	3/40 (7.5)	[b]	0.25[c]
Maternal or child serious adverse event during pregnancy and birth	70/610 (11.5)	59/616 (9.6)	0.83 (0.58-1.16)	0.27
Childhood (2 years of age)				
Health				
Composite outcome of death or moderate/severe neurodevelopmental impairment at 2 years	51/419 (12.1)	67/399 (16.8)	1.45 (0.98, 2.15)	0.064

Moderate/severe neurodevelopmental impairment	35 /403 (8.7)	47/379 (12.4)	1.48 (0.98,2.33)	0.087
Individual components of disability				
Motor	4/456 (0.9)	4/461 (0.9)	0.99[d] (0.25,3.98)	0.99
Cognitive function	18/452 (4.0)	19/461 (4.1)	1.03 (0.58, 1.84)	0.92
Hearing	2/465 (0.4)	1/466 (0.2)	0.56 [d] (0.33,0.94)	0.028
Speech and language	14/446 (3.1)	18/445 (4.0)	1.32 (0.72,2.43)	0.36
Vision	4/466 (0.9)	0/447 (0.0)	[b]	0.13[c]
Respiratory	3/434 (0.7)	7/413 (1.7)	3.03 [d] (1.56,5.88)	0.0011
Gastrointestinal	4/432 (0.9)	9/412 (2.2)	2.67 [d] (1.37,5.20)	0.004
Renal	1/434 (0.2)	3/414 (0.7)	3.65 (1.96,6.82)	<0.0001
Admitted to hospital during follow up	51/434 (11.8)	48/416 (11.5)	0.98 (0.65, 1.47)	0.92
Behavioral scale scores at 2 years assessed in strengths and difficulties questionnaire[c]			OR of abnormal score	
Total difficulties	9.8 ± 4.9, n=302	10.2 ± 4.9, n=295	1.23 (0.85,1.78) [f]	0.28
Emotional problems	1.1 ± 1.2, n=341	1.1 ± 1.2, n=328	1.01 (0.61, 1.67) [f]	0.96

Conduct problems	2.7 ± 1.8, n=342	2.6 ± 1.8, n=326	0.92 (0.65, 1.31)[f]	0.66
Hyperactivity scale	4.2 ± 2.4, n=334	4.5 ± 2.3, n=315	1.10 (0.79, 1.55)[f]	0.57
Peer problems scale	2.0 ± 1.7, n=345	2.1 ± 1.6, n=318	1.22 (0.88, 1.69)[f]	0.22
Prosocial scale	6.3 ± 2.2, n=339	5.9 ± 2.3, n=320	1.20 (0.88, 1.63)[f]	0.25
Impact scale	0.2 ± 1.0, n=424	0.2 ± 1.2, n=404	1.31 (0.73,2.35)	0.37
EQ-5D				
Change in EQ-5D from baseline to birth	-0.023 ± 0.220, n=199	-0.021 ± 0.207, n=191	0.001 (-0.034, 0.036)	0.97
Change in EQ-5D from baseline to 12 months	-0.015 ± 0.221, n=274	-0.009 ± 0.213, n=279	0.003 (0.026,0.032)	0.83
Women's views				
Women's perception of treatment 1m post delivery – proportion extremely or fairly satisfied	314/327 (96.0)	294/307 (95.6)	0.93 (0.42,2.04)	0.85

Binary outcomes are summarised as number of women with outcome / number of women with information on outcome (percentage with outcome of those with information). Continuous outcomes are summarised as mean ± SD, number with information on outcome.

[a] Median (range) of time to death (days) of 759 (1, 1331) and 751 (1, 1335) respectively

[b] Regression failed with and without adjusting for previous pregnancy of > 14 weeks

[c] Exact fisher test

[d] Not adjusted for previous pregnancy of at least 14 weeks because regression failed

[e] Mean ± SD age in weeks at assessment of 107.7 ± 17.7 (placebo) and 106.9 ± 17.1 progesterone

[f]Score analysed as binary variable (raised compared with normal score)

Table 5 – Safety outcomes

	Placebo	**Progesterone**
N	590	593
Pregnancy complications	**n / denominator (%) or mean** ± SD	
Maternal		
Obstetric cholestasis	6 (1.0)[a]	4 (0.7)
Hypertension	24 (4.1)	23 (3.9)
Pre-eclampsia	11 (1.9)	10 (1.7)
Eclampsia	1 (0.2)	0 (0.0)
Preterm premature membrane rupture	72 (12.2)	65 (11.0)
Antepartum haemorrhage	36 (6.1)	37 (6.2)
Gestational diabetes	37 (6.3)	27 (4.6)
Confirmed DVT	2 (0.3)	0
Cervical cerclage	39 (10.8)[b]	41 (11.1)[c]
Other maternal complication	164 (27.8)	166 (28.0)
Fetal		

Any	18 (3.1)	19 (3.2)
of which		
AC < 5th centile	4 (22.2)	6 (31.6)
Liquor volume reduced	6 (33.3)	6 (31.6)
Doppler > 95th centile (umbilical artery)	1 (5.6)	1 (5.3)
Absent EDF (umbilical artery)	0 (0.0)	1 (5.3)
Reversed EDF (umbilical artery)	1 (5.6)	1 (5.3)
Abnormal antenatal CTG	7 (38.9)	3 (15.8)
Hospital admissions		
Antenatal hospital admissions per woman, mean ± SD and median [range]	0.7 ± 1.3 [d] 0.0 [0, 10]	0.6 ± 1.1 [e] 0.0 [0, 8]
Women with hospital admission for reasons other than PTL	135 (23.2)[d]	107 (18.5)[e]
Women with hospital admissions for TPL	132[d]	119 [e]
with tocolysis	18 (3)	15 (3)

with steroid	71 (12)	80 (14)
with antibiotic	52 (9)	38 (7)
with cervical cerclage	10 (2)	8 (1)
with magnesium sulphate	0 (0.0)	0 (0.0)
Labour		
Duration of first stage in hours	4.1 ± 5.1, n= 463	4.3 ± 5.3, n= 470
Duration of second stage in minutes	47.0 ± 132.8, n= 462	41.2 ± 91.6, n=471
Duration of third stage in minutes	17.0 ± 46.2, n= 465	16.1 ± 51.6, n= 477
ARM required	131 /468 (28.0)	122/448 (27.2)
Analgesia in labour (any)	455 /576 (79.0)	478/574 (83.3)
of which		
General anesthetic	16 (2.7)	12 (2.0)
Epidural	191 (33)	197 (34)
Opiates	88 (14.9)	88 (14.8)
Nitrous oxide	269 (47)	303 (53)
Other	34 (5.8)	31 (5.2)
Delivery method[f]		

Spontaneous Vaginal Delivery (SVD)	380 (65.7)	375 (65.1)
LSCS in labour	58 (10.0)	57 (9.9)
LSCS pre-labour	92 (15.9)	84 (14.6)
Forceps	21 (3.6)	27 (4.7)
Ventouse	18 (3.1)	20 (3.5)
Vaginal breech (spontaneous or assisted)	9 (1.6%)	13 (2.3)
Blood loss, ml, mean (SD)	387.4 ± 356.4, n=572	423.7 ± 393.8, n=572
Blood transfusion	10 /578 (1.7)	18 /574 (3.1)
Antibiotics during labour and delivery	96 /578 (16.6)	92/573 (16.1)
Surgical procedure required	15 / 578 (2.6)	18 / 575 (3.1)
Duration of hospital stay in days	3.2 ± 2.2, n=577	3.3 ± 4.1, n=567
Duration of hospital stay in days, median [range]	3.0 [1.0, 19.0], n=577	3.0 [1.0, 86.0], n=567
Any postpartum complication	83 /580 (14.3)	90/577 (15.6)
Placental examination, N available	84	83
No evidence of infection	57 (67.9)	56 (67.5)
Chorioamnionitis	10 (11.9)	9 (10.8)
Chorioamnionitis and funisitis	17 (20.2)	18 (21.7)

Birth outcomes		
Male sex	289 / 578 (50.0)	293 /578 (50.7)
Birthweight in g	2822 ± 884, n=577	2875 ± 847, n=577
Apgar score at 1 minute - Median [IQR]	9.0 [8.0, 9.0], n= 553	9.0 [8.0, 9.0], n= 557
Apgar score at 5 minutes - Median [IQR]	9.0 [9.0, 10.0], n= 555	9.0 [9.0, 10.0], n= 560
Length of hospital stay days - Median [IQR]	2.0 [1.0, 6.0], n= 556	2.0 [1.0, 4.0], n= 562
Outcomes at 2 years		
Weight in kg	13.2 ± 2.6, n = 355	13.4 ± 2.7, n =332
Height in cm	87.2 ±10.7, n=369	87.4 ± 7.9, n=347
Head circumference in cm	48.9 ± 4.6, n=354	49.6 ± 6.7, n=332

Outcomes in the safety population – women who took at least one tablet of placebo or progesterone. Outcomes available for the number of women at the top of the column, unless otherwise stated.

[a] Available for 589 women.

[b] Available for 360 women.

[c] Available for 368 women.

[d] Available for 581 women.

[e] Available for 579 women.

[f] Available for 578 in the placebo and 576 in the progesterone group respectively.

Figure 1. Consort diagram

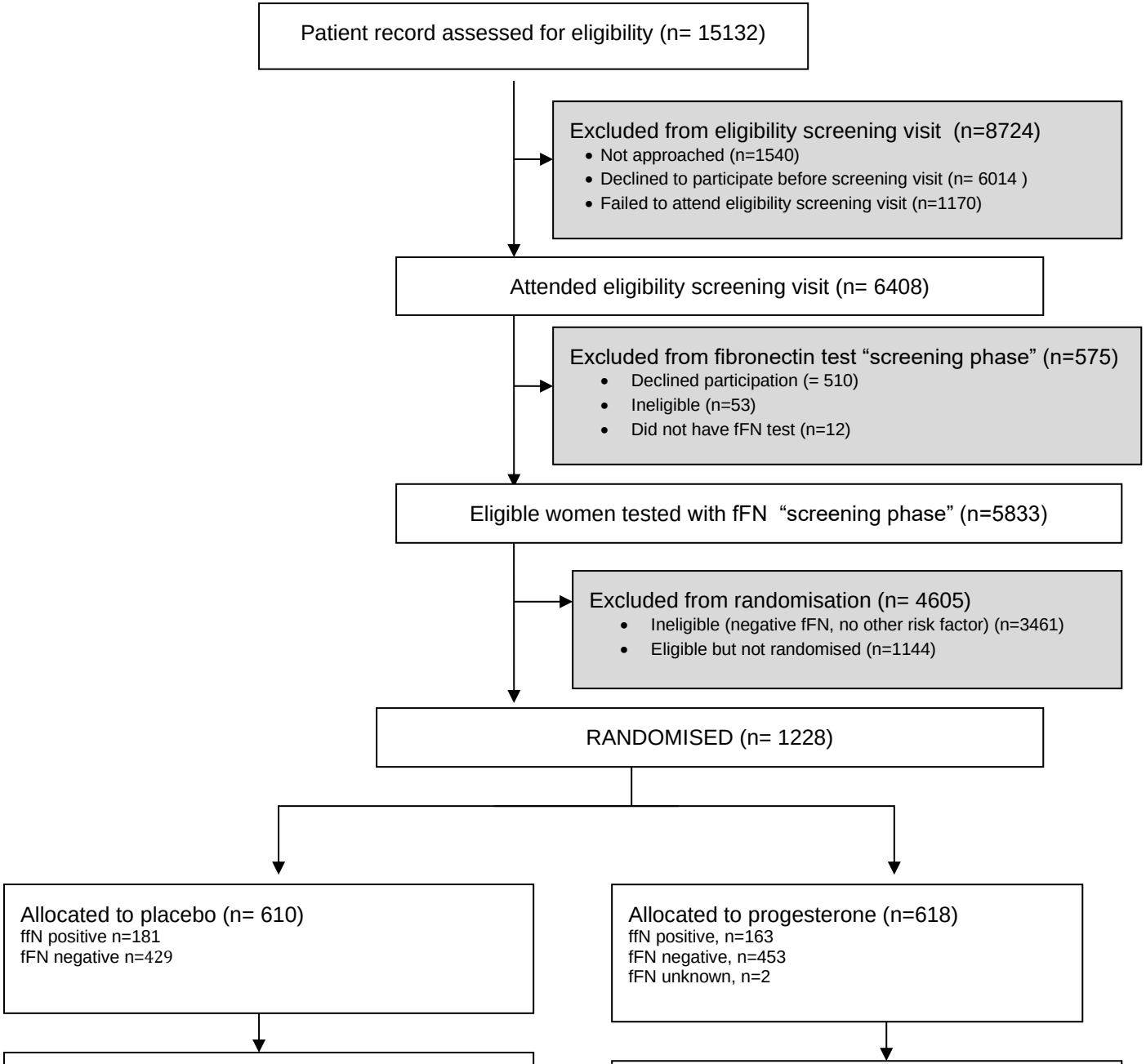

Figure 1 Legend

a. Randomised in error – ineligible for treatment, excluded post randomisation.

[b] Note consent withdrawals for each of the phases refer to consent withdrawal at any time prior to reaching the outcome for that phase

c Note losses to follow up for each of the phases refer to losses to follow up at any time prior to reaching the outcome for that phase